DOI:
10.1016/j.comnet.2018.01.030

A Hybrid DMM Solution and Trade-off Analysis for Future Wireless Networks

J. Carmona-Murillo[a,*], V. Friderikos[b], J. L. González-Sánchez[c]

[a]*Departament of Computing and Telematics Engineering, University of Extremadura, Spain*
[b]*Centre for Telecommunications Research, King's College London, UK*
[c]*Research, Technological Innovation and Supercomputing Center of Extremadura (CénitS), Spain*

Abstract

Mobile Internet data traffic has experienced an exponential growth over the last few years due to the rise of demanding multimedia content applications and the increasing number of smart mobile devices. Seamless mobility support at the network level is envisioned as a key architectural requirement to deal with the ever increasing demand for data and content, cell densification and to efficiently utilize a plethora of heterogeneous wireless access networks (HetNets). Current and emerging efforts on that frontier aim to evolve mobility management protocols towards a more distributed operation to tackle shortcomings that stem from fully centralized oriented approaches. However, as will be detailed hereafter, there are instances where distributed mobility management result in lower performance, which might affect real time and several over the top (OTT) applications (as well as incur increased levels of signaling overhead in the network). To this end, in this paper we provide a meticulous analysis of the different trade-offs between centralized and Distributed Mobility Management (DMM) and based on the analysis we propose a Hybrid DMM solution that overcomes, in terms of mobility costs, both centralized and distributed mobility management protocols. Furthermore, we also conduct a comprehensive analytic and numerical comparison of the different mobility solutions. Our results indicate the significant benefits in terms of packet delivery cost and signaling overhead that Hybrid DMM solutions might bring. Finally, we conclude by discussing some open ended issues in mobility management in emerging and future wireless networks.

Keywords: IPv6 mobility management, hybrid solution, cost analysis, performance evaluation.

1. Introduction

Over the last few years, IP-based mobility management in the Internet has been one of the most active research fields in communications. Mobility management protocols are responsible for maintaining the ongoing communications while the user roams among distinct networks (changing points of attachment) and also to provide reachability for mobile users in such heterogeneous environment in terms of access. The existing IP mobility support protocols developed by the IETF (Internet Engineering Task Force) are all based on centralized mobility anchors that manage the traffic and signaling of the Mobile Nodes (MNs). The two most representative centralized mobility management protocols are Mobile IPv6 (MIPv6) [1] and Proxy Mobile IPv6 (PMIPv6) [2]. MIPv6 introduces a home agent (HA) as a mobility anchor, while PMIPv6 tries to manage mobility locally (i.e., at the foreign network) by introducing a centralized agent called local mobility anchor (LMA). This node is responsible for both mobility signaling and user data forwarding. However, centralized mobility management protocols need to be redesigned in order to cope with the recent trends in mobile Internet and the current increasing mobile data traffic demand. This demand is expected to continue rising with an almost exponential trend even for the foreseeable future [3].

*Corresponding author
Email addresses: jcarmur@unex.es (J. Carmona-Murillo), vasilis.friderikos@kcl.ac.uk (V. Friderikos), joseluis.gonzalez@cenits.es (J. L. González-Sánchez)

Moreover, as mobile data traffic increases, the growth in signaling load is expected to increase almost 50% faster than the growth in data traffic over the next few years [1]. The generated amount of control information is increasing dramatically for Evolved Packet Core (EPC), and is expected to grow even more with the deployment of Long Term Evolution Advanced (LTE-A), as the access network is connected directly to base stations, managing all signaling traffic. Looking further into the future, 5G networks that will entail inevitably cell densification (i.e., smaller cells) to increase overall network capacity will result in an even increased signaling cost for managing mobility. One of the keys to signaling traffic explosive growth is the increasing number of Internet-connected mobile machine-to-machine devices and applications with high mobility demands that result in heavy control data. These requirements in both data and growing signaling traffic demand has become a critical consideration for network operators when dimensioning and planning mobile networks to meet a satisfactory user experience. In this increasingly heterogeneous and complex environment, efficient mobility management can be deemed as a key functionality related to the overall network performance, due to its implication in control and data planes.

In order to address these limitations which inherently occur in Centralized Mobility Management (CMM) protocols, Distributed Mobility Management (DMM) solutions are being developed to efficiently handle the current mobile traffic explosion. In DMM, the core idea is that the mobility anchors are distributed within the network, topologically closer to the users, with the aim to provide an almost optimal routing support and an efficient use of network resources to improve the scalability required for next generation mobile networks [4].

However, and as already alluded to above, despite the fact that a number of mobility management approaches are on-design phase towards a more distributed operation aiming to mitigate the problems related to centralized operation, there are instances where DMM incurs higher costs and the performance of the network might be compromised. In fact in some of these cases, CMM seems to solve more efficiently the mobility problem and therefore should be preferred. Particularly those in which cell resident time is short and/or the number of remaining active sessions in previous networks is high. These situations happen, for example, when an MN moves frequently and it begins new sessions in different visited networks. In these cases, the performance of DMM approaches fall down due to the number of tunnels that need to be managed by the distributed nodes.

As stated in [5], future mobile network architectures might potentially exhibit a hybrid behavior in which the mobility management of some traffic will be kept centralized, while mobility support for other applications can be distributed. Network virtualization and software defined networking techniques that would allow flexible and programmable networks based at the control and data user plane will allow to efficiently utilize hybrid mobility schemes as the ones proposed hereafter.

In this paper, a Hybrid DMM mobility management scheme is proposed, that adapts to the specific topological characteristics of the infrastructure network of mobile operators, in which the data and signaling traffic are forwarded following a centralized or distributed scheme depending on, hereafter detailed, decision criteria for protocol selection. The key benefit of the proposed hybrid solution is that it manages to reduce significantly both signaling and routing cost. To the best of our knowledge, this is the first effort to exhibit a hybrid centralized-distributed solution for future mobile network architectures.

The rest of the paper is organized as follows. In Section 2, we briefly present closely related work in the area of mobility management. Section 3 details the background and the motivations, highlighting the evolution of the IP mobility management and introducing the benefits of hybrid solutions; and details of such hybrid scheme are described in Section 4. Section 5 defines the network model and system parameters. Section 6 introduces the decision criteria algorithms in which are based the hybrid solutions. The cost analysis are presented in Section 7. Section 8 shows the numerical evaluation. Finally, concluding remarks from this work are given in section 9.

[1]Signaling is growing 50% faster than data traffic, 2012 `http://nsn.com/index.php?q=system/files/document/signaling_whitepaper_online_version_final.pdf`

2. Related work

During the last few years, mobility management at the IP level attracted significant attention from both industry and academia, and it has been an active field in communications research. This has been mainly driven by the increased heterogeneity of wireless access which calls for solutions at the IP level in order to support session continuation when mobile users change their point of attachment. Relevant standards development organizations such as IETF and 3GPP (Third Generation Partnership Project) are making ongoing efforts to address the new needs in mobile IP networks; these works have recently resulted in some proposals to create an evolved architecture of the current mobile networks [6] [7]. Current packet-based mobile architectures, such as 3GPP EPS (Evolved Packet System) and WiMAX make use of IP as the enabling technology for both voice and data communications. Therefore, IP mobility management protocols will inevitably play a key role to address continuity and session persistence throughout user movement among different networks. At the same time mobility control at the IP layer has been considered a network management tool for provisioning load balancing and/or data offloading in heterogeneous wireless networks [8].

The main IP mobility management proposals are based on MIPv6 and PMIPv6. Fig. 1 shows an overview of both protocols [9].

In order to enable the mobility service in MIPv6, the Mobile Node (MN) is assigned with a permanent home address in its Home Network (HN), and establishes a connection with the communication peer, the Correspondent Node (CN). A Home Agent (HA) serves as the anchor node in the HN that tracks the network connection point (location) of a user as the user moves. Periodically, or whenever the user changes their point of attachment to the network, the user registers with the HA through Binding Update (BU) messages, informing of its current location and establishing a tunnel between the HA and the MN located in a visited network. In MIPv6, the HA is the centralized part of the system since it is on the critical path of both signaling and data for mobile users.

Mobility in Mobile IPv6-based solutions requires the host to send mobility management signaling messages to the home agent, which is potentially located -topologically- far from the visiting network. In addition to performance issues for supporting seamless session continuity this means that the protocol requires stack modification of the mobile node in order to support the mobility improvements. In addition, the requirement for the modification of mobile nodes may cause them to become increasingly complex.

Network-based protocols on the other hand, are mainly derived from PMIPv6. PMIPv6 is based on MIPv6 in the sense that it extends MIPv6 signaling and reuses many concepts such as for example the HA functionality. The new principal functional entities of PMIPv6 are the mobile access gateway (MAG) and local mobility anchor (LMA). The MAG typically runs on the AR. Its main role is to detect MN's movements and initiate mobility-related signaling with the MN'LMA on behalf of the MN. In addition, the MAG establishes a tunnel with the LMA to enable the MN to use an address from its home network prefix and emulates the MN's home network on the access network for

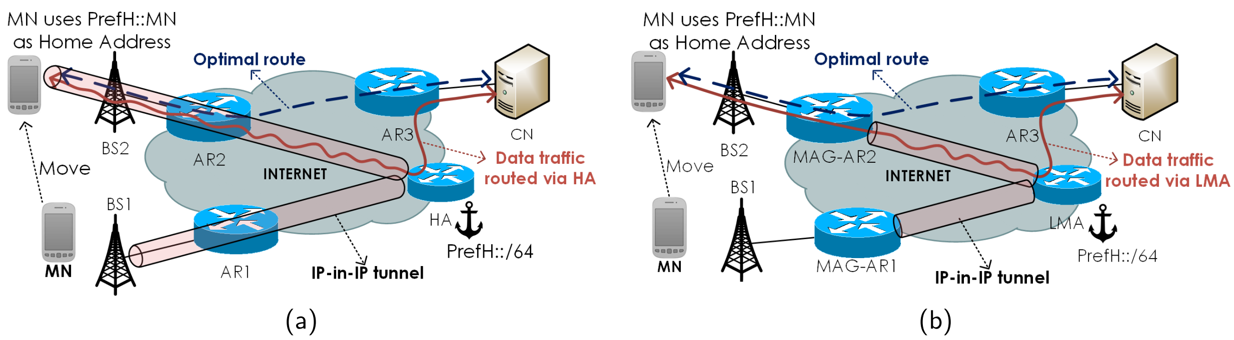

Figure 1: Centralized Mobility Management solutions. (a) MIPv6 and (b) PMIPv6.

each MN. On the other hand, the LMA is similar to the HA in MIPv6. Its main role is to maintain access to the MN's address while it moves and to store information necessary to associate an MN with its serving MAG, enabling the relationship between the MAG and LMA to be maintained.

3. Background and motivation

3.1. CMM limitations

As briefly alluded to above, CMM requires a single handed mobility anchor, e.g., HA at MIPv6 and LMA at PMIPv6 to allow for session continuity when MNs are moving across different networks. Nowadays, most of the deployed architectures have a small number of centralized anchors managing the traffic of thousands of mobile users. According to current MAG specifications, anchor devices support a maximum of 64,000 attached MNs [10]. These centralized approaches have certain limitations for handling a large volume of mobile data traffic such as non-optimal routing, scalability problems and reliability/robustness. These limitations have been identified in [11] and summarized below,

- Non-optimal routing: In CMM, all traffic is routed through a central mobility anchor, resulting in a longer path and thereby, increasing the end-to-end delay.
- Scalability problems: With the increase of mobile nodes and the traffic explosion of mobile data, the centralized anchor needs to be able to deal with an increasing aggregate volume of traffic. Here arise scalability problems due to the mobility signaling processing and routing resources that the mobility anchor needs to manage for such increased amount of data traffic.
- Reliability: Centralized anchors are more vulnerable to single point of failures and attacks than a distributed system.

3.2. The evolution from Centralized to Distributed Mobility Management

In order to address the above mentioned limitations of centralized mobility management solutions, a new paradigm has been recently proposed which gained attention: the so-called Distributed Mobility Management. In essence, DMM develops a new concept for handling mobility, with the main characteristic being that the mobility anchors are placed closer (topologically) to the user, distributing the control and data plane mobility functions among entities located at different places on the core/access network.

Depending on the role of the mobile node in the handover process, mobility management protocols can be broadly classified in two categories, namely those that require active involvement of the MN and those that do not. For example PMIPv6 does not require the MN to be involved in the layer 3 signaling to complete a handover.

A representative proposal of a DMM solution which is based on Mobile IP is detailed in [12] and [13] (Host-Based DMM, HB-DMM). In this work, the authors attempt to improve the performance of mobility support by distributing mobility agents (called AMA) at the edge of the access network and the MN is served by a mobility anchor located in the serving network. When a MN moves to an adjacent network, a tunnel is created between the serving AMA and the origin AMA, located in its Home Network and a new address is configured in the MN. As depicted in Fig. 2(a), this solution creates multiple tunnels between AMAs and in cases where a high mobility rate exists, the system performance might be critically compromised by the frequent registrations and maintenance of multiple tunnels.

Network-Based DMM (NB-DMM) [14] in turn, exempt the MN from participation in any mobility signaling and therefore there is no need for network software upgrade at the MN for mobility support since distributed mobility anchors perform mobility signaling on behalf of the MN, as is the case for example in PMIPv6. In this proposal, like the one previously mentioned, the mobility management functionalities are moved to the Access Routers (ARs) level in order to anchor the traffic closer to the MN. Each AR is required to have both mobility anchoring and location functionalities, and it is referred to as mobility capable access router (MAR). In NB-DMM, a new session is anchored at the current AR and initiated using the current IPv6 address. When a handover occurs before the end of the session, the data traffic is tunnelled between the current MAR and the anchoring MAR for the session. The basic operation of NB-DMM is depicted in Fig. 2(b).

3.3. Towards a flexible network: Hybrid solutions

The evolution from CMM to DMM approaches has shown clear signs of achieving better utilization

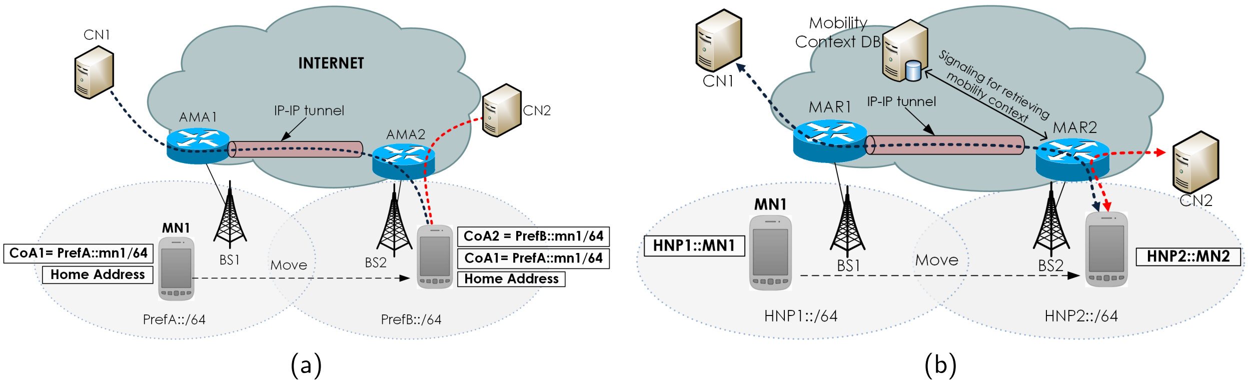

Figure 2: Distributed Mobility Management solutions. (a) Host-Based DMM and (b) Network-Based DMM.

of resources in the network, outperforming the traditional protocols and optimize mobility management performance [13], [14], [15], [16]. Despite these facts, there are some scenarios in which the DMM paradigm also incurs high delivery costs and possible significant signaling overheads [5].

In the DMM operation, each distributed MAR/AMA that manages a MN keeps a bidirectional tunnel between itself and all MAR/AMA where a session of the MN was originated and is still alive. This situation can be deemed as inefficient in certain circumstances. In these cases, a CMM based behavior might be preferred.

Additionally, next generation mobile networks will be driven by software, relying in emerging technologies such as software-defined networking (SDN) and network function virtualization (NFV) [17]. These technologies will make the network more flexible as a new functionality can be introduced with simple software upgrades, and more sophisticated algorithms can be employed to manage the network. In particular, a hybrid CMM-DMM mobility management solution, might provide the benefits of each one of them resulting in better overall network performance. In this respect, a virtualized programmable network infrastructure provide the flexibility of locating the mobility anchor in different locations based on the underlying topology.

In the next section the proposed network-based hybrid mobility management scheme is detailed and a set of decision criteria to determine how the traffic can be more efficiently anchored are discussed.

4. Description of the hybrid mobility management scheme

The proposed hybrid solution can be deemed as an amalgamation of previous schemes, where mobility management in an IPv6 network can be handled by a centralized protocol such as PMIPv6 or by a distributed one such as NB-DMM. One key motivation of our hybrid solution is to take advantage of some aspects from PMIPv6 and others from NB-DMM, minimizing the limitations of centralized and distributed mechanisms, developing a solution that allows the network to decide, depending on network performance based criteria, when to manage the traffic in a distributed way or when to keep it centralized. It has to be mentioned that although in this paper the hybrid mechanism relies on network-based mobility management solutions, it could be equally applied to host-based environments.

4.1. Initial mobility anchoring

At an initial state of the network operation, each AR is selected to manage the traffic with one of the two possible protocols (PMIPv6 or NB-DMM) although they can offer support in both mechanisms, if needed. This means that the operators can benefit from this flexible anchoring, since future networks will be based on NFV and SDN. This way, mobility function can be virtualized and the anchoring can be programmable in order to efficiently manage the network resources. In the next sections we discuss, in detail, the decision criteria used in this approach.

When a mobile node attaches to a Base Station (BS), the AR in charge of the BS assigned for the new session provides the MN an IP address ensuring IP reachability and/or IP session continuity, as well as the respective routing/forwarding support from the establishment of the session.

As we can see in Fig. 3, when MN1 joins the access subnet, the responsible AR (AR1) retrieves the IP address for the MN following the registration procedure of PMIPv6 so, AR1 will act as MAG and the LMA will be the anchor point for the MN. When MN2, located in a different subnet, attaches to a BS, it initiates the NB-DMM registration procedure. In this case, AR5 will act as MAR so, AR5 will be the origin anchor router for the MN2.

While MNs remains attached to the same AR, all connections will be managed with the initial mobility management protocol. This initial mobility anchoring is transparent to the user.

4.2. Registration and data delivery mechanisms

This section describes the hybrid solution operation when a mobile node moves among several points of attachment associated to different mobility management protocols, and the forwarding data mechanisms. Once the MNs are initially anchored to a centralized or distributed agent (LMA or MAR respectively), they can deliver data with their correspondent nodes following the PMIPv6 or NB-DMM operation. No changes are needed to the initial attachment or the initial forwarding mechanism.

Fig. 4 shows the MN's movement from the access network of AR1 to the access network of AR5 and how new sessions are managed. In Fig. 4(a), MN1 is initially registered to AR1 and a session with CN1 is active. Upon an IP handover from one MAG (AR1) to another (AR2) as it is shown in Fig. 4(b),

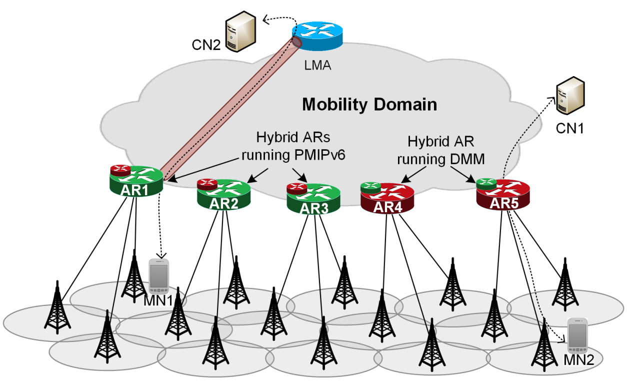

Figure 3: Overview of hybrid approach.

the registration procedures starts. The binding is updated at the LMA as follows. The previous MAG (AR1) sends a proxy binding update (PBU) to the LMA deregistering the MN, and the new MAG (AR2) sends a PBU to the LMA registering it. The LMA replies to each by a proxy binding acknowledgement (PBA). Regarding the data plane of PMIPv6, when a new session arrives (session 2 in this case) from CN2, the centralized sessions of MN1 are anchored to the LMA and all the MN data traffic passes through the LMA. A tunnel is established for this purpose between the LMA and the serving MAG (AR2).

When the MN1, managed by centralized anchors, moves to a DMM AR (AR3) as is shown in Fig. 4(c), if any centralized connection remains active, the AR registers them with the PMIPv6 procedure. This is possible because all ARs in the hybrid approach are both PMIPv6 and DMM capable. If during its stay in AR3 the MN establish a new connection with CN3 (session 3), it will be anchored following the DMM procedure and allowing centralized and distributed connections at the same time. For the DMM signaling, the protocol relies on a database (DB) that stores ongoing mobility sessions for the MNs. Specifically, it stores the home network prefixes currently allocated to the MN and their respective anchoring points. Finally, in Fig. 4(d), the MN moves to a new DMM AR (AR4). In this case, the new serving AR will act as a MAR and performs the handover management following the regular operation of NB-DMM as follows. The new MAR (AR4) retrieves the IP address of the anchoring MAR (AR3) from the database and sends a PBU to each anchoring MAR, in this specific case it would be AR3. Then, AR3 replies by a PBA and a bidirectional tunnel is created among AR4 and each anchoring MAR, in this case between AR4 and AR3. If a new session arrives (session4), the new traffic is routed directly between the MN and the correspondent node (CR4). The old traffic (session 3) is routed through the anchoring MAR (AR3) using the previously established tunnel.

5. Network model

A communication network can be defined as a directed graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes the set of nodes (vertices) and $\mathcal{E}$ denotes the set of links (edges) interconnecting the nodes. Let $M \subseteq \mathcal{V}$ be the set of routers that serve as the Mobility Anchors

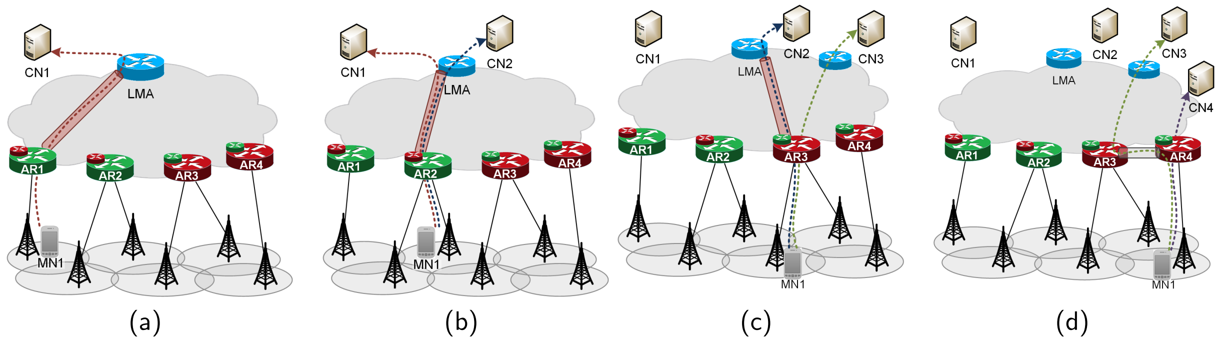

Figure 4: Data path during hybrid solution operation. (a) Initial State. MN1 is attached to AR1. Session 1 is created. (b) MN1 moves to AR2 and session 2 is created. (c) MN1 moves to AR3. Session 1 finishes and session 3 arrives. (d) MN1 moves to AR4. Session 2 finishes and session 4 arrives.

for the mobile nodes, $\mathcal{K} \subseteq \mathcal{V}$ be the set of access routers in the network and $\mathcal{N}$ the set of mobile nodes moving around in the network. Each node n_i $(1 \leq i \leq |\mathcal{N}|)$ is equipped with network interfaces that enables them to be reachable through wireless technologies. We further assume a number of base stations belonging to the set $\mathcal{B}$ that provide full coverage in the scope of the geographical area under consideration.

In this scenario, the access routers (AR) are the first hop router and, it is the MN's point of attachment at the radio and the IP levels. The CNs, without loss of generality, are assumed to be stationary for simplicity. We denote by h_{x-y} the average hop distance, i.e. average number of hops, between two network entities x and y. The average hop distance is assumed to be symmetric, i.e. $h_{y-x} = h_{x-y}$.

The handover probability between adjacent ARs is assumed to be known from previous historical data and measurements and denoted by the $\mathcal{H}$ matrix, as follows,

$$\mathcal{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1k-1} & h_{1k} \\ h_{21} & \ddots & & & h_{2k} \\ \vdots & & \ddots & & \vdots \\ h_{k-11} & & & \ddots & h_{k-1k} \\ h_{k1} & h_{k2} & \cdots & h_{kk-1} & h_{kk} \end{pmatrix}$$

Each element of this matrix h_{ij} is the probability of handover occurring between ARs i and j. As mentioned above the handover probability matrix can be obtained for a given network topology from historical network traces and statistics and are normally known to a mobile operator.

5.1. Traffic Model and User Mobility

We consider a scenario where a MN might be actively engaged simultaneously with several CNs in the Internet, i.e., having several active sessions. Without loss of generality we assume that the sessions from a MN are generated to follow a Poisson process with mean rate λ_s (i.e. the inter-arrival time between sessions is exponentially distributed with this rate). We assume also that the duration of a typical session is exponentially distributed with mean rate μ_s. By modeling the scenario as a system under the probability distribution of a typical $M/M/\infty$ queue, the probability of having at least one session at a time is equal to λ_s/μ_s.

As with respect to the mobility model, we assume the Random Waypoint (RWP) model, since it is one of the most widely used mobility models in the literature due to its simplicity, realistic behavior and ease of implementation [18]. RWP model is in essence a synthetic model that describes the movement pattern of independent MNs on a finite continuous plane. In RWP, a mobile node moves from one waypoint to the next waypoint by randomly choosing its destination coordinates, its speed of movement, and the amount of time that it will pause when it reaches the destination. On reaching the destination, the node pauses for some time distributed (Θ) according to some random variable and the process repeats itself. Once the pause time expires, the node chooses a new destination, speed, and pause time. In this paper, an one-dimensional line segment $[0,\phi]$ is considered to calculate the expected distance between one waypoint to the next waypoint $E(L)$. According to [19], $E(L)$ is as follows:

$$E(L) = \frac{1}{3}\phi \tag{1}$$

If the velocity of an MN v is constant and $v > 0$ during its entire movement process, then the expected transition time $E(T)$ is:

$$E(T) = \frac{1}{v}E(L) \tag{2}$$

Let $E(C)$ denote the number of subnet crossings during the transition. By using the previous equations, we can estimate the average residence time $E(R)$ of the MN in a subnet as follows:

$$E(R) = \frac{E(T) + \Theta}{E(C)} \tag{3}$$

Using as a starting point the model presented in this section and in order to demonstrate that a Hybrid DMM outperforms the overall routing cost than previously proposed DMM solutions, we obtain the following,

Lemma 1. *Let $c_{p,q}$ be the routing cost to forward packets from node p to node q in the network. Let $C^H_{i,j}$ and $C^D_{i,j}$ be the total routing cost of the network, for whatever two nodes i and j, using the Hybrid DMM mechanism and using the NB-DMM mechanism respectively, then $C^H_{i,j} \leq C^D_{i,j}$.*

Proof. Let $\mathcal{K}$ denote the set of access routers in the network and M be the set of routers that serve as the centralized mobility anchors (LMA) for the mobile nodes. Let $g \in \mathcal{M}$, i and $j \in \mathcal{K}$, i and j be two adjacent routers. Since $c_{g,i} \leq c_{g,j} + c_{i,j}$ because all paths are SP (Shortest Path), then Lemma 1 holds, as required. □

6. Decision criteria

As we have eluded to in previous sections, the ARs in the hybrid approach are both PMIPv6 and DMM capable. However, they are initially selected to operate in a centralized or distributed way depending on some criteria that the operator can select. In this section we introduce two decision criteria algorithms that carry out this protocol selection. These algorithms make decisions based on network information. As explained previously, DMM provides some clear benefits with respect to CMM, but the performance of DMM is very topology dependant [13] since the flows have to be tunnelled from the old AR to the new AR. This means that, in some cases, depending on the routing path between the two ARs are involved in the handover phase CMM should be preferred instead of DMM. For this reason, the main criteria that is used in our algorithms is the topology information.

As it is shown in Fig. 5, the first algorithm uses only information about the network topology, whereas the second one uses both topology and location information of the BSs. In the sequel we detail two algorithms based on these criteria aiming to optimize network performance and provide specific benefits to network operators by allowing various degrees of freedom in deciding mobility management procedures.

6.1. Node-assignment algorithm

In Algorithm 1, which is called hereafter the node-assignment algorithm, the decision is made according to the actual operation of the mobility management protocols. The data plane of distributed protocols extend the data path during the movement of the MN through the ARs at the edge of the network, whereas centralized protocols anchor all sessions of a MN to the same entity, the mobility anchor.

Hence, the algorithm in essence declare for each candidate AR if they will be acting as a mobility anchor or not, which will as a result define the degree of mobility function decentralization in the network.

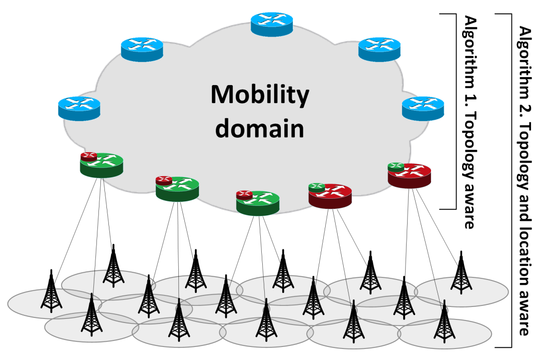

Figure 5: Level of information in order to handle user mobility at the different ARs.

Algorithm 1: Node-assignment decision criteria

Input: Adjacency matrix $AdjMatrix$
Output: Array $protocolAR$ with the protocol to use in each AR

foreach $AR\ i \in K$ **do**
  $V_i = ARNeighbours(i)$;
  foreach $AR\ j \in V_i,\ j \neq i$ **do**
    $routingCost(i,j)$=dijkstra(i,j,AdjMatrix);
  end
end
foreach $AR\ i \in K$ **do**
  $\alpha(i) = \frac{\sum_{j=1}^{V_i} routingCost(i,j)}{|V_i|}$
end
$\overline{\alpha} = mean(routingCost)$;
foreach $AR\ i \in K$ **do**
  if $\alpha(i) > \overline{\alpha} + threshold$ **then**
    $protocolAR(i) = PMIPv6$;
  else
    $protocolAR(i) = NB - DMM$;
  end
end
return $protocolAR$;

Fig. 6 shows an example of the decision criteria process that, as a result, select the optimal mobility anchoring that the AR (Node 2) should use according to the topological information. In this example, Node 2 has three neighbouring ARs, so we consider the routing cost to connect with each one of them. Let $c_{i,j}$ be the cost to reach node j from node i, for the Node 2 we consider $c_{2,1}$, $c_{2,3}$ and $c_{2,4}$. With these three costs, we obtain α_2, which represents the mean routing cost for Node2 to reach any of its AR neighbours. The algorithm checks if this aggregate cost is above or below the average value ($\overline{\alpha}$) of all the α_i. There can be a number of different policies about the threshold cost value that can be used in order to take into account current network conditions and/or provide a weight for the different mobility anchoring options.

If the value α_i is larger than ($\overline{\alpha}$), it means that the network is sparse in that area and it would be preferable to use a centralized protocol. On the other hand, a lower value of α_i means that the AR is well connected to its neighbours and a distributed approach is a more efficient option.

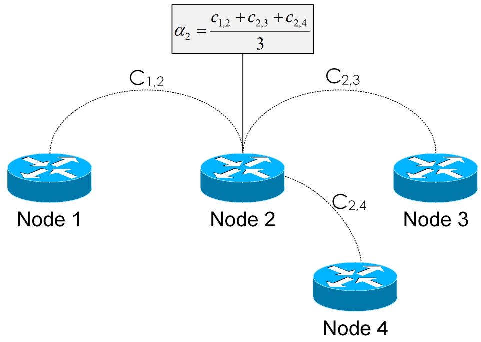

Figure 6: Example of anchoring selection by Node 2 using Algorithm 1.

6.2. Link-assignment algorithm

The previous algorithm establishes a heuristic mechanism to select the mobility management protocol to be used at each AR, using topological information from the network. As we have previously detailed, the way of selecting a centralized or distributed protocol for each AR is based on the routing cost to its neighbours given by a link state algorithm. Intuitively speaking, this means that the base stations located in a well connected area of the network will be assigned to operate with the DMM protocol, whereas base stations in topological areas of the network which are not well connected will be assigned to operate in a centralized way. However, using only the topological information may mean that some sessions are not managed properly, especially when the value of α_i is not representing correctly the routing cost $c_{i,j}$ values, i.e, one of the values $c_{i,j}$ is significantly higher and the rest of the cost $c_{i,k}$ is small (in other words there is high deviation from the calculated mean cost that is used in the node-assignment algorithm). In those cases the movement of the MNs between the ARs i and j may be enhanced by differentiating between the use of appropriate mobility anchoring selection. For this reason, we propose a second algorithm that, instead of deciding the protocol that will be used per node (at each AR), selects mobility anchoring per link. Using this link-assignment algorithm we can obtain a more efficient decision, avoiding the problems that may be encountered by the first algorithm, at the extra cost of using more detailed network topology information.

The main idea behind this second algorithm (see Algorithm 2) is to make the decision closer to the

Algorithm 2: Link-assignment decision criteria

Input: Adjacency matrix $AdjMatrix$

Output: Matrix $mLinks$ with the protocol to use at each link

foreach $AR\ i \in K$ **do**
  $V_i = ARNeighbours(i)$;
  foreach $AR\ j \in V_i,\ j \neq i$ **do**
    $routingCost(i, j)$=dijkstra(i,j,AdjMatrix);
  end
end
$(K, X) = buildOverlayARGraph(K, routingCost)$;
$\overline{c} = mean(routingCost)$;
foreach $Link\ l_{i,j}\ in\ X$ **do**
  if $c_{i,j} > \overline{c} + threshold$ **then**
    $mLinks(i, j) = PMIPv6$;
  else
    $mLinks(i, j) = NB - DMM$;
  end
end
return $mLinks$;

mobile node, taking into account the area through which the MN is moving. Thus, the decision considers the link cost between the current AR the MN is attached to, and the AR to which the MN is moving.

In order to formally define an algorithm that takes into account the above mentioned information, we make use of an overlay network, where the nodes are the ARs and the links are the connection between them following the next model. Starting from a network architecture where each AR manage several BSs, we define an Overlay AR Graph (see Fig. 7) obtained with the following process. Let $G = (\mathcal{V}, \mathcal{E})$ be the graph that defines the network as presented in Section 5, and G' the overlay graph, $G' = (K, X)$, where K is the set of ARs in the network defined by G, $K \subseteq \mathcal{V}$, each $AR \in K$ cover a set of BSs defined by the operator. X is the set of links, $x_{i,j}$ denote a link $(0, c_{i,j})$ integer variable that is set to $c_{i,j}$ if there is any BS managed by an Access Router $AR_i \in K$ adjacent to any other BS managed by $AR_j \in K$, otherwise there is no link between i and j. In this case, and without loss of generality, $c_{i,j}$ represents the routing cost from i to j which can be obtained with a shortest path algorithm over the graph G.

Thus, a link is created in the overlay graph between two ARs, if the MN can move directly from one BS managed by the first AR to another BS managed by the second AR. This overlay graph consider only those links that allow the direct movement of an MN between two ARs, improving the decision criteria.

Finally, to identify the mobility anchoring scheme that will be assigned to each link, we calculate the average value of all routing costs, defined as $\overline{c}$. If the value of the cost $c_{i,j}$ is greater than $\overline{c}$, the AR will use PMIPv6 as the mobility management protocol in the connections associated to the link between AR i and AR j, and NB-DMM otherwise.

With the overlay AR graph G', we are able to identify the related location information needed to make decisions in the proposed hybrid mobility management mechanism. The underlay graph G corresponds to the physical network topology, and the overlay level is a logical one, and contains the connections and associated costs between adjacent ARs.

The degree of each node in the overlay network G' indicates the number of connections with adjacent ARs. Each one of these links relates to a mobility anchoring assignment that will be used to determine how the AR that manage that link handle the mobility of the MNs arriving and leaving the associated BSs.

7. Analytical evaluation

When a MN moves and is about to change its point of attachment, signaling is initiated between the network entities involved in that mobility event and a change in the routing path is needed in order to deliver the data packets to the new point of attachment of the MN. Moreover, the mobility protocols using tunnelling mechanisms and the associated overhead by such encapsulation need to be taken into account for each scheme. In this section we evaluate the parameters involved in mobility aspects, such as the cost functions of registration updates (C_u), traffic routing (C_r) and tunnelling overhead (C_t). We first define a quantitative analysis of these three main characteristics followed by numerical investigations to provide a more in-depth analysis.

For the purpose of offering a more complete understanding, we first compare the proposed Hybrid DMM approach with both host-based and network-based distributed mobility management (HB-DMM, NB-DMM) with the centralized

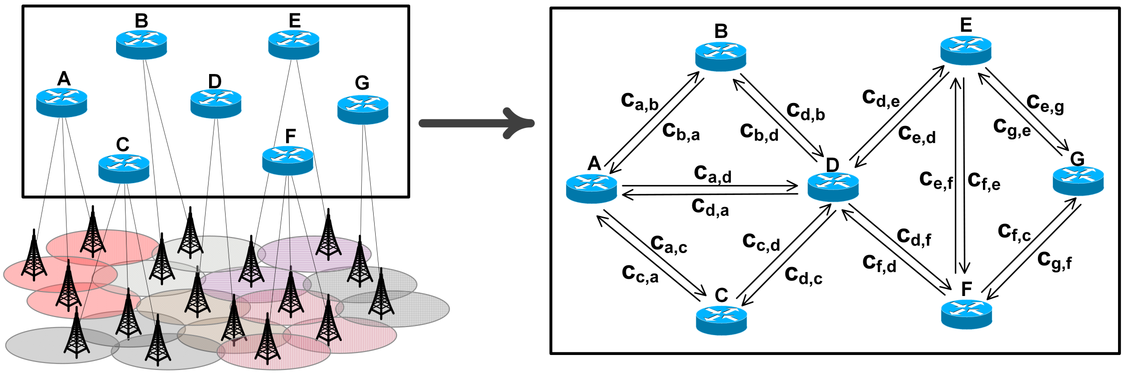

Figure 7: Physical AR and BS structure and Overlay AR graph.

mobility management solutions; namely MIPv6 and PMIPv6. Table 1 shows the parameters description used in this analysis.

7.1. Signaling cost

One of the main functionalities for any IP mobility management protocol is the process of maintaining the MN's mobility session up to date. Such a task requires control messages that needs to be sent among the mobility agents in the network. In this section we refer to the signaling cost of registration update for a session as C_u. This value reflects the traffic load when signaling messages are sent, i.e., this cost depends on the size of the signaling messages and the number of hops in every level 3 handover process during the time interval that the MN communication remains active. Therefore, considering the handover probability h_{ij} that a node moves between two adjacent networks i and j, and the average number of handovers in a session $E(R)/t_s$, described in section 5, the signaling cost (C_u) is defined by the actual size in bytes of the signaling messages multiplied by the number of required hops as follows:

$$C_u = \sum_{i=1}^{k} \sum_{j=1, j \neq i}^{k} h_{ij} \cdot \frac{E(R)}{t_s} \cdot S_u \tag{4}$$

In each case, C_u represents the signaling cost specifically for a protocol and is described next. A registration is performed for each new movement between neighbouring subnets. In Mobile IPv6 and PMIPv6, the registration update is needed with the centralized mobility anchor (HA or LMA respectively). The distributed approaches, such as Host-Based DMM and Network-Based DMM, update their movements with the distributed anchors. In HB-DMM, the mobility anchor is called AMA and is located in the access router so the AMA is the first IP capable router for the MN. In this approach, when a mobile node moves, the MN registers its movement to the serving AMA (sAMA) and it establishes bidirectional tunnels with previous AMAs the MN was connected to. A similar process occurs with NB-DMM, where upon a handover, the new MAR retrieves the IP addresses of the anchoring MARs for the MN from the DB. The new MAR then registers the MN at all these MARs. In the case of distributed mobility protocols, we define n as the number of valid addresses configured at the MN in HBDMM and NB-DMM ($n =$ number of handovers $+ 1$).

Regarding the proposed hybrid proposal, the average probability that the sessions are managed by centralized or distributed protocols are denoted by $\mathcal{P}_C$ and $\mathcal{P}_D$ respectively. Moreover $C_u(Hybrid)$ consists of the sum of the cost of the sessions that

Table 1: Parameter description

Param.	Description
t_s	Average connection time for a session
s_d	Average size of a data packet
s_u	Av. size of a control message for update
s_t	IPv6 tunnel header size
λ	Transmission rate for a downlink packet
h_{ij}	Handover probability from i to j
$E(R)$	Average residence time
h_{x-y}	Number of hops between x and y

are managed in a distributed manner and those that remain centralized.

$$\mathcal{P}_C + \mathcal{P}_D = 1 \tag{5}$$

Hence, we obtain the following values for the signaling cost when the registration update process occurs:

$$C_u(MIPv6) = 2 \cdot s_u \cdot h_{MN-HA} \tag{6}$$
$$C_u(PMIPv6) = 2 \cdot s_u \cdot h_{MAG-LMA} \tag{7}$$
$$C_u(HBDMM) = [2 \cdot s_u \cdot h_{MN-sAMA} + + \sum_{i=1}^{n-1}(2 \cdot s_u \cdot h_{AMA_i-sAMA})] \tag{8}$$
$$C_u(NBDMM) = [2 \cdot s_u \cdot h_{AR-DB} + + \sum_{i=1}^{n-1}(2 \cdot s_u \cdot h_{AMA_i-sAMA})] \tag{9}$$
$$C_u(Hybrid) = \mathcal{P}_C(C_u(PMIPv6)) + + \mathcal{P}_D(C_u(NBDMM)) \tag{10}$$

7.2. Data packet delivery cost

One of the key challenges that mobile networks have to deal with is the increasing amount of data traffic mainly generated by high quality multimedia and video/audio streaming sessions. The introduction of distributed solutions aims to reduce this cost by avoiding all user data to traverse a centralized anchor. However, DMM might not be adequate in certain situations due to their topology dependent performance. As we described in Section 3, distributed mechanisms creates multiple tunnels between the anchors and in cases where a high mobility rate exists, the system performance might be critically compromised by the frequent registrations and maintenance of multiple tunnels.

With the analysis developed in this section, in conjunction with the numerical results obtained from the simulations, we evaluate and measure the network load in terms of total data packet delivery cost for a session. This metric is defined as C_r and its value is controlled by the size of the data messages scaled by the number of hops needed to forward packets from the CN to the MN or vice versa.

In MIPv6 and PMIPv6, packets are routed from the CN to the MN's anchor (HA or LMA respectively), and forwarded from the MN's anchor to the MN. In HB-DMM and NB-DMM, when a MN moves, the traffic established in the new network will be routed directly to the CN whereas the remaining connections will be tunneled to its corresponding anchoring MAR and then routed to the CN. With our hybrid DMM proposal, mobility is managed by PMIPv6 or NB-DMM depending on the decision criteria described in Section 6. Hence, the packet delivery cost values for each solution are as follows:

$$C_r(MIPv6) = \lambda \cdot [(s_d \cdot h_{CN-HA} + + (s_t + s_d) \cdot h_{HA-MN})] \tag{11}$$
$$C_r(PMIPv6) = \lambda \cdot (s_d \cdot h_{CN-LMA} + + (s_t + s_d) \cdot h_{LMA-MAG} + + s_d \cdot h_{MAG-MN}) \tag{12}$$
$$C_r(HBDMM) = C_r(NBDMM) = = \lambda \cdot (C_d^r + C_i^r) \tag{13}$$
$$C_r(Hybrid) = \mathcal{P}_C(C_r(PMIPv6)) + + \mathcal{P}_D(C_r(NBDMM)) \tag{14}$$

where, C_d^r and C_i^r are the units of cost of delivering one packet in the direct and indirect modes of DMM, respectively [14]. Then, these costs are expressed as follows

$$C_d^p = s_d \cdot h_{CN-sMAR} + s_d \cdot h_{sMAR-MN} \tag{15}$$
$$C_i^p = s_d \cdot h_{CN-MAR} + (s_t + s_d) \cdot h_{MAR-sMAR} + + s_d \cdot h_{sMAR-MN} \tag{16}$$

7.3. Tunnelling Cost

To achieve seamless mobility support, mobility management protocols use a tunnel to forward/re-direct packets. Depending on the operation of each proposal, a certain quantity of those packets will be tunnelled and, therefore, the delivery cost will be penalized with the tunnelling overhead. In that respect, MIPv6 and PMIPv6 for example, tunnel data flows from a centralized anchor to the MN or the MAG agent. On the other hand, in distributed based solutions data flows are tunnelled between the decentralized anchors. Although DMM solutions mitigate the tunnel overhead, in certain circumstances when the MN is running long-lasting applications or the residence time is short, tunnelling management becomes a critical metric in network performance due to the overhead that it introduces. Similarly to the signaling cost and the packet delivery cost, our hybrid DMM proposal adapts its operation to the network topology due to the decisions taken at the beginning of the network operation.

The tunnelling cost metric represents in essence the cost of adding a tunnelling overhead to the overall data packet delivery cost. So, the tunnelling cost can be derived from packet delivery cost by setting the payload size of the packet to zero, $s_d = 0$.

$$C_t(MIPv6) = \lambda \cdot s_t \cdot h_{HA-MN} \tag{17}$$

$$C_t(PMIPv6) = \lambda \cdot s_t \cdot h_{LMA-MAG} \tag{18}$$

$$C_t(HBDMM) = \lambda \cdot \sum_{i=1}^{n-1} s_t \cdot h_{AMA_i-sAMA} \tag{19}$$

$$C_t(NBDMM) = \lambda \cdot \sum_{i=1}^{n-1} s_t \cdot h_{MAR_i-sMAR} \tag{20}$$

$$C_t(Hybrid) = \mathcal{P}_C(C_t(PMIPv6)) + + \mathcal{P}_D(C_t(NBDMM)) \tag{21}$$

8. Performance evaluation

Following the analysis presented above, this section aims to provide insights about the impact of several mobility costs on the overall network performance, as well as evaluate the assessment of the algorithms under different topologies. Thus, firstly, we compare the introduced hybrid CMM-DMM solution in its two versions, node-assignment algorithm (Hybrid Solution 1) and link-assignment algorithm (Hybrid Solution 2), with MIPv6 and PMIPv6 as centralized approaches, as well as HB-DMM and NB-DMM as distributed protocols in terms of registration update cost, packet delivery cost and tunnelling overhead. Secondly, the evaluation under different topologies is carried out.

8.1. Evaluation of mobility costs

The evaluation through simulations aims to study the distributed approaches in a more realistic environment. The platform selected for the evaluation through simulations was MATLAB. The scenario defined for the evaluation is illustrated in Fig. 8. Due to the dependence on the topology of DMM protocols, we selected this asymmetric topology due its mixture between a well connected hierarchical network and a sparse network. This will produce more reliable results because the nodes will move around the connected and the sparse areas of the network, avoiding misleading performance of centralized or distributed protocols due to the network topology. In addition such topology will allow us to shed further light on the dependency on network topology on the performance of different mobility management protocols.

The traffic and mobility parameters values used in the simulations, as well as the numerical results of mobility costs are presented next.

We consider a scenario where an MN may traverse several simultaneous active sessions with several CNs in the Internet. We assume that session arrivals to a MN follows a homogeneous Poisson process with mean rate $\lambda_s = 0.01$ (i.e. the inter-arrival time between sessions is exponentially distributed with this rate). We assume also that the duration of a typical session is exponentially distributed with mean session duration $\mu_s = 10$ time units [20].

We validate our model using a Random Waypoint mobility model as presented in Section 5.1 with the following parameters. Speed: uniformly distributed between 1 and 10 m/s; Pause interval: uniformly distributed between 1 and 5 min.; Walk interval: uniformly distributed between 5 and 20 min [16]. In order to drive the evaluation towards a more realistic scenario, we also run the simulations with real-world mobility track logs obtained from users carrying GPS receivers. The sample settings where traces are obtained are two university campuses (one in Asia and one in the US), one metropolitan area (New York City), one State fair and one theme

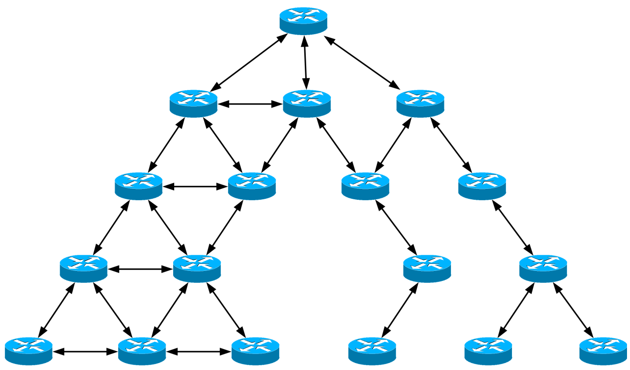

Figure 8: Topology used in the simulation.

park (Disney World). The participants walk most of the time and may also occasionally travel by bus, trolley, car, or subway trains. These settings are selected because they are conducive to collecting GPS readings [21].

Moreover, the simulation is run for a different number of MNs, ranging from 1 to 50 and the simulation time is sufficiently large (45000 units) to avoid "typical runs" statistical problems. The dimensions of the simulation scenario for the RWP mobility model is a rectangular area of 5x5 km^2 and the MNs are initially located randomly in that area. With regards to the real mobility tracks, the dimensions of the rectangular simulation area is set to be the same as in the GPS traces. In all simulation scenarios, we used the same initial positions found in the respective real traces for the same number of users.

In the evaluation, the simulations are repeated 25 times to improve the accuracy of the results with a confidence interval of 95%.

Fig. 9 shows the accumulated signaling cost of registration delay update vs. the number of MNs during all the simulation execution. Both mobility models are shown in order to compare human mobility as shown in Fig. 9(a), and the Random Waypoint results shown in Fig. 9(b).

In this case, in MIPv6 and PMIPv6 the control messages are exchanged between two entities, the serving AR (MAG in PMIPv6) and the centralized mobility anchor (HA in MIPv6 and LMA in PMIPv6). In these centralized solutions, all sessions are anchored to the same agent so, all of them are updated in the same signaling message, introducing a low overhead. On the other hand, in HB-DMM and NB-DMM, the control messages are exchanged between the distributed nodes for each connection that remains active for MNs during their movement; thus, there can be scalability concerns since the signaling overhead increases in proportion

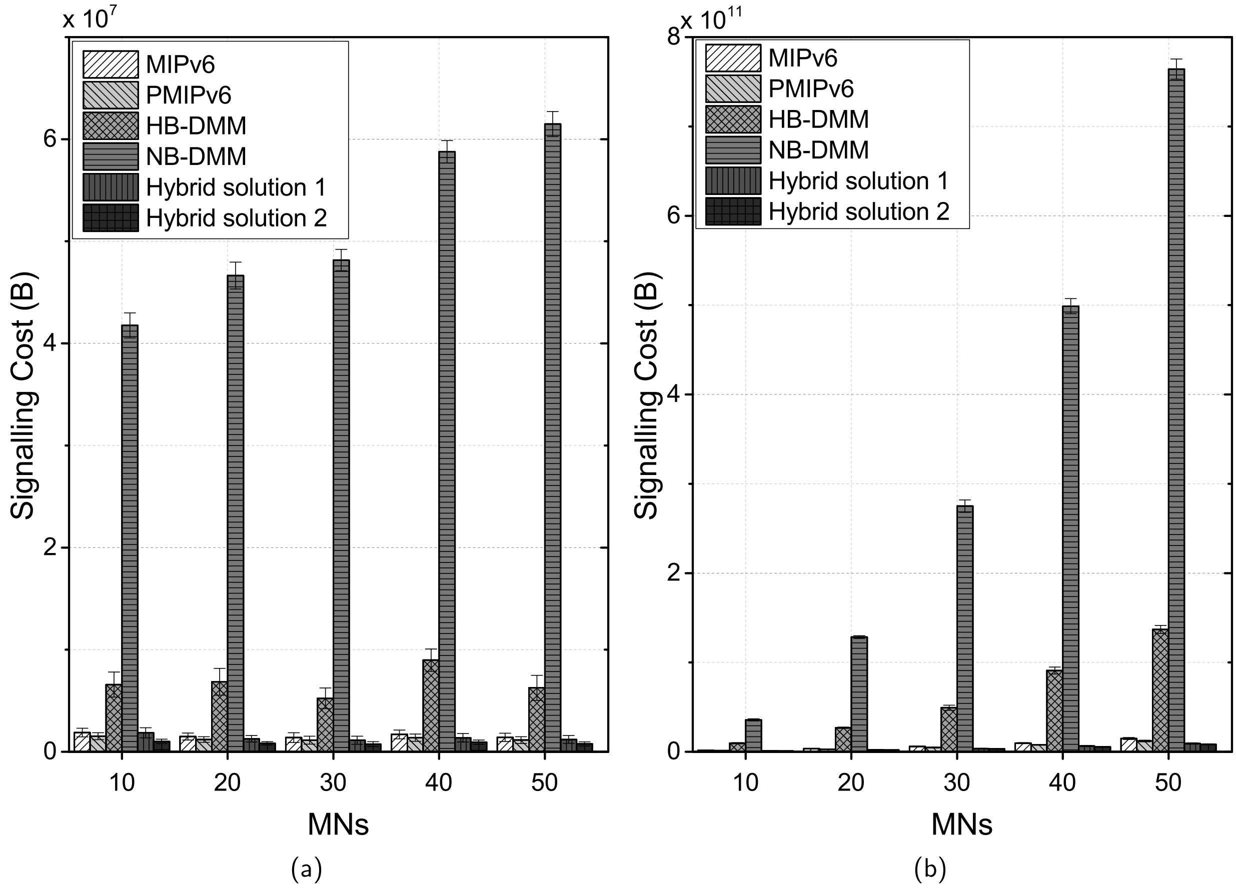

Figure 9: Signaling Cost. (a) Human walk mobility model; (b) Random Waypoint mobility model

to the number of the MN's IP prefixes/addresses anchored at ARs other than the serving AR. With respect to the proposed hybrid mobility proposals, both algorithms manage the signaling overhead in an efficient way, offering an equivalent performance to centralized approaches. As is shown in Fig. 9, the signaling cost in distributed approaches is significantly higher than the other protocols, especially NB-DMM. This high cost in NB-DMM is produced because it additionally requires a control message to be exchanged with the database each time a handover is produced, adding an extra signaling overhead.

With respect to the mobility models, the RWP mobility demonstrates a clear trend as the number of MNs increases. This effect is not visible when using the realistic human based mobility traces because when some users are added to the scenario, they do not add any additional movement so, the signaling cost of registration updates is not increased, whereas in the RWP mobility model, all users behave more similarly from a mobility point of view.

With respect to the accumulated packet delivery cost, the results obtained are shown in Fig. 10. It can be observed that both mobility models present similar behavior because the packet delivery cost is not highly dependent on the mobility model. In the same figure it can also be observed how centralized solutions perform non-optimal routing and therefore the overall cost is higher. DMM protocols on the other hand, outperform the cost of centralized protocols although when a handover occurs, packets are routed through a suboptimal path. As we can see, the value of both DMM protocols is the same. This occurs because the data plane in both HB-DMM and NB-DMM operates in the same manner.

Based on the above observation we can conclude from Fig. 10 that the benefits obtained from

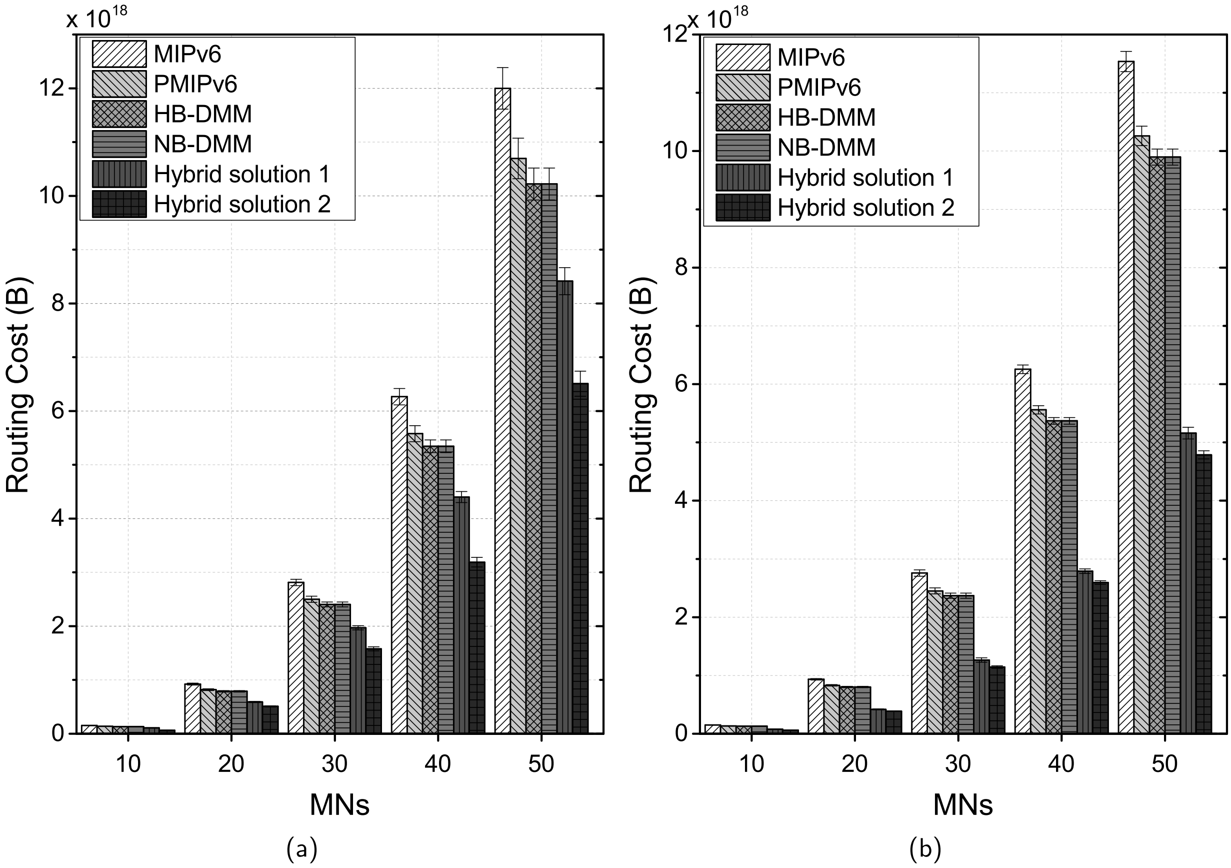

Figure 10: Routing cost. (a) Human walk mobility model; (b) Random Waypoint mobility model

the deployment of hybrid solutions in mobile networks can produce substantial improvements, up to 60% reduction in routing cost with respect to centralized approaches and up to 45% gains with respect to DMM protocols. Taking into account the increasing traffic expected for future mobile networks, this saving in resources due to the inclusion of the proposed hybrid approaches can facilitate the deployment of next generation mobile network architectures.

Finally, in Fig. 11 the tunnelling cost of the mobility protocols is compared. The significant difference between CMM and DMM solutions is highlighted. While HB-DMM and NB-DMM introduce insignificant tunnelling, centralized solutions cause a high overhead in the network due to the tunnelling process. With the proposed hybrid approaches, we try to maintain all mobility sessions with good performance in terms of signaling and packet delivery cost. Although the tunnelling cost of Hybrid DMM improves significantly the cost of CMM protocols, they introduce more tunnelling overhead than HB-DMM and NB-DMM. This effect is produced because the algorithms proposed for protocol selection in our Hybrid DMM take the decisions based on routing cost, because it is the metric with higher impact in a mobile network. Thus, and taking in account the low tunnelling introduced by DMM approaches, when some sessions in Hybrid DMM are managed by PMIPv6 in order to reduce the overall routing cost, the tunnelling cost is affected negatively. This little penalization can not be avoided and is acceptable from a global performance perspective. Fig. 11 illustrates that the tunnelling overhead introduced by Hybrid DMM solution is minimal.

It is worth to note that, despite the fact that the signaling data can be seen as insignificat with respect to routing or tunnelling cost, signaling has become a critical when dimensioning packet core network. The growing year by year of signaling data requires that the design of mobile networks

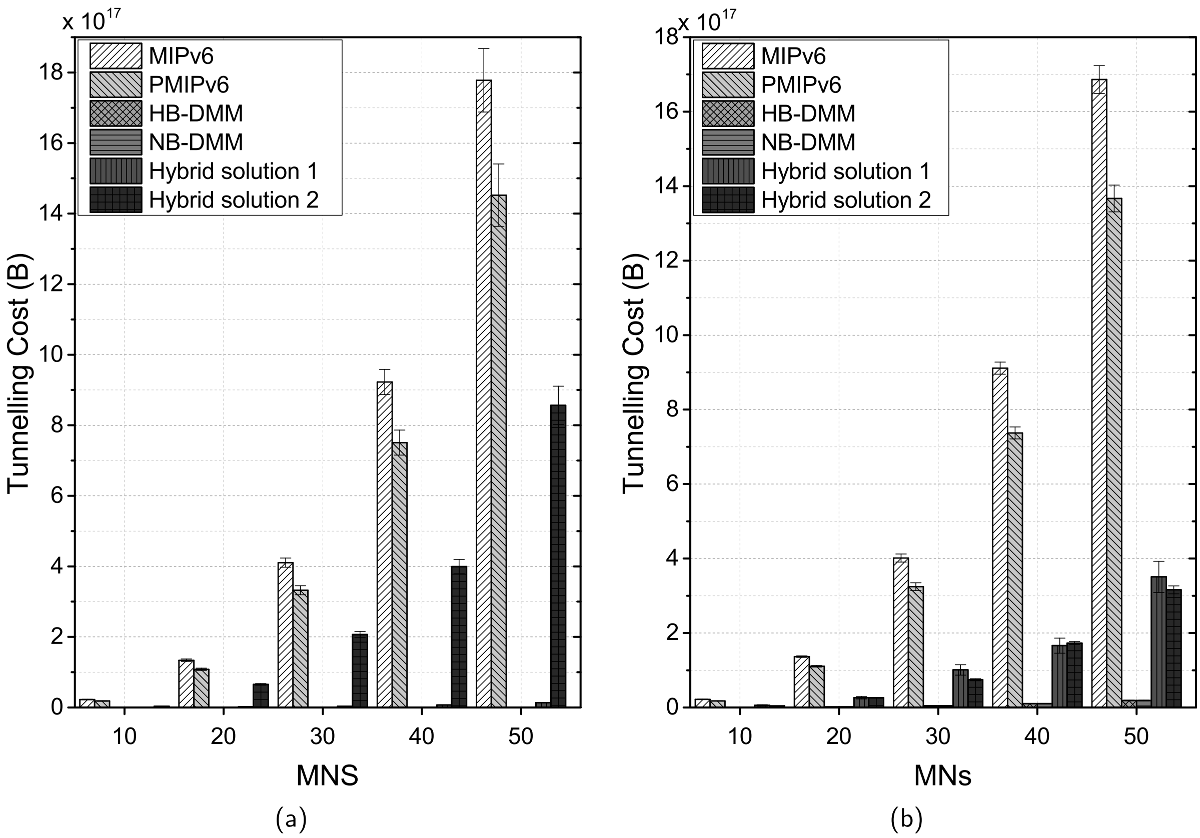

Figure 11: Tunnelling cost. (a) Human walk mobility model; (b) Random Waypoint mobility model

take into account this issue. As it is stated in [22], this is one challenge that operators must contend with. Considering that low latency is one of the main requirements for 5G networks, the reduction of signaling required to set up and maintain a data connection, can reduce the E2E latency.

8.2. Evaluation of the decision criteria algorithms with different topologies

In order to assess the sensitivity of the algorithms and to carry out a fairer evaluation of them, this section aims to provide insights about the impact of the topology in the proposed hybrid mechanisms. Therefore, apart from the balanced topology used in the previous section, the simulations have been conducted under two additional topologies, shown in Fig. 12.

These topologies have been selected due to its connectivity degree. Fig 12.a is a low-connected topology in which nodes are not connected with other nodes in the same level (siblings nodes), whereas Fig. 12.b is given by a full-connected graph in which nodes have a connection with all their parents and siblings.

Thus, the results shown in this section have been carried out with the same simulation parameters than in the previous section. In order to evaluate the impact of the topology, three different topologies have been considered: Balanced topology (Fig. 8), low-connected topology (Fig. 12.a) and full-connected topology (Fig. 12.b). In this case, the MNs move following the RWP model.

The results of the mobility costs in low-connected and full-connected networks are shown in Fig. 13 and Fig. 14 respectively. Comparing these data with the evaluation of the metrics with the balanced topology (Fig. 9.b, Fig. 10.b, and Fig. 11.b), some interesting results can be highlighted.

The first thing that it is necessary to take into account is how the Hybrid solutions selects, for each AR, the mobility management protocol that manage the sessions in each topology. In general, the ARs located in well connected areas of the network are assigned to operate with NB-DMM protocol, whereas ARs in topological areas of the network which are not well connected are assigned to operate in a centralized way. According to this, the low connected topology manage most of the sessions in a centralized way, whereas in the full connected topology, the preferred protocol will be NB-DMM.

A comparison of the signaling cost in these topologies shows that this metric achieve a better behavior in a full-connected topology. This is because the edge nodes are connected directly and the tunnels among the serving MAR and the other MARs from which there is still an active connection are short in terms of number of hops. This connection allows a shorter tunnel length during the movement of the MN, reducing the signaling cost. On the contrary, this metric achieve the highest values in low-connected topologies.

The routing cost is the metric that exhibit a larger variation among the results of Hybrid solutions in the three topologies. This is due to the fact that the algorithms are based on this cost to take the decisions. As it can be seen in the results, the highest improvements achieved with the hybrid solutions are those in which there are possibilities of selecting the better paradigm that a new connection must follow. This happens in the balanced topology and, to a lesser extent, in the low-connected network. On the other hand, the behavior of hybrid

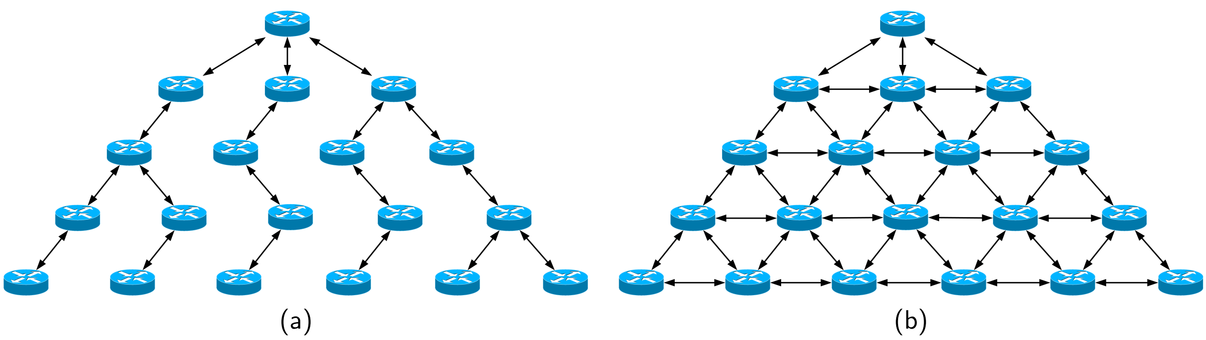

Figure 12: Topologies used in this evaluation. (a) Low-connected topology; (b) Full-connected topology

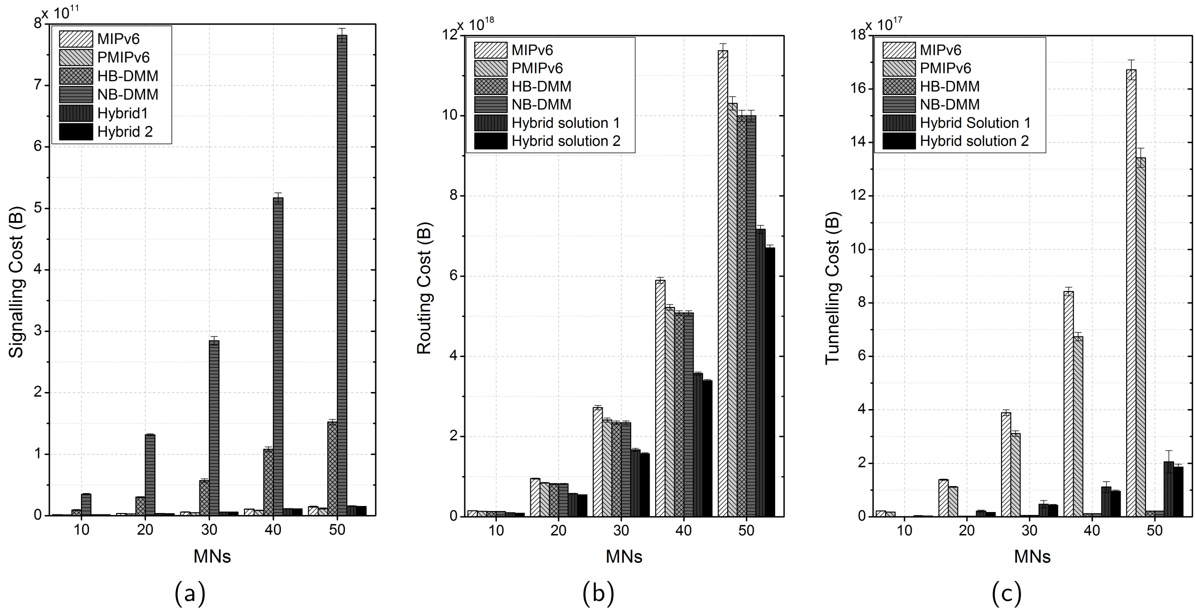

Figure 13: Evaluation of the low-connected topology (a) Signaling cost; (b) Routing cost; (c) Tunnelling cost

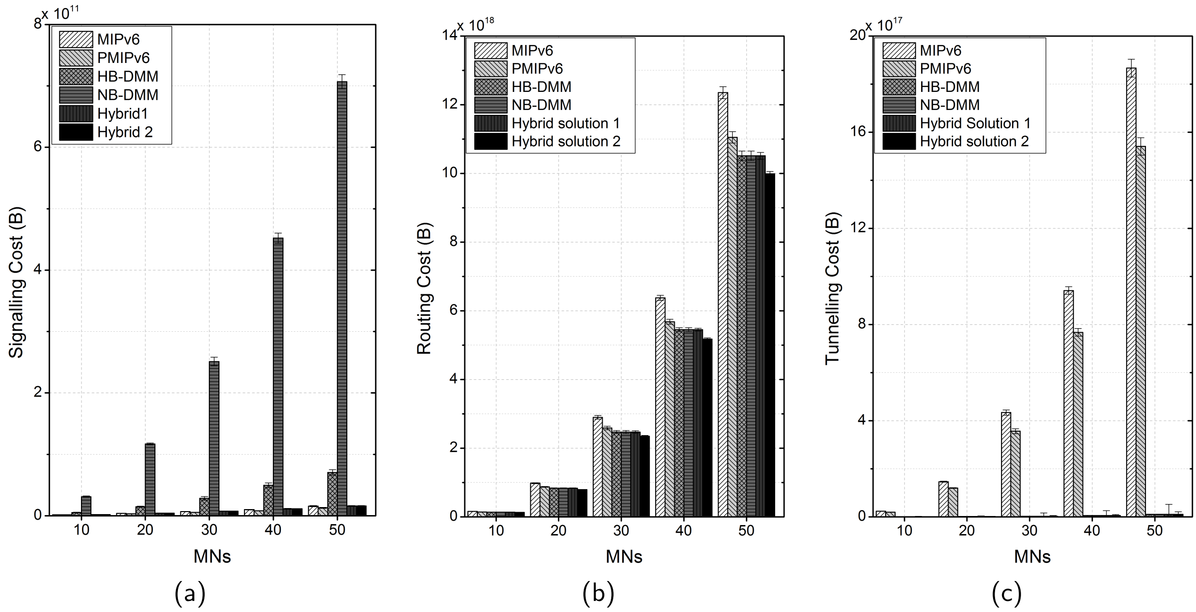

Figure 14: Evaluation of the full-connected topology (a) Signaling cost; (b) Routing cost; (c) Tunnelling cost

solutions in full-connected topologies is similar to DMM protocols. This is because, in this kind of topologies, and according to the first decision criteria algorithm, all connections are managed in a distributed way. The second decision criteria achieves almost the same values but with a minor improvement due extra information used to perform the selection.

If we focus on the data plane metrics, i.e., routing cost and tunnelling cost, the results show how the reduction in the routing cost brings a penalty in the tunnelling cost. As can be observed in the results, the balanced topology achieves the best results in the routing cost compared with the other topologies. However, the tunnelling cost is higher than the results obtained with low-connected and full-connected topologies. This is the expected result, as the goal of the decision criteria algorithms is to optimize the forwarding plane of the network.

We can conclude that the Hybrid mechanisms achieve the best results in topologies with different degrees of connectivity, in which the algorithms can take decisions and the network metrics can be improved selecting CMM or DMM suitably, minimizing its limitations.

9. Conclusions

Mobility management protocols are evolving towards a distributed operation in order to deal with the steadily increasing mobile Internet traffic. In essence, distributed mobility management solutions rely on distributed anchors where mobility management functions are located. With distributed operation, the support of mobility results in significant less overhead in terms of encapsulation, while routing, scalability and robustness issues are also improved. However, as we have detailed, in some scenarios, the performance of DMM can decrease dramatically depending on the shortest path between ARs as well as a number of additional factors, such as the high session arrivals rate, long and frequent movements (i.e, short residence time) and long-lasting sessions. In these situations the operation of DMM might lead to a lower performance and hence the use of a centralized mobility management solution would be a preferred option.

Based on the above mentioned observations, this paper proposed a Hybrid DMM solution in which mobility can be managed by a centralized protocol such as PMIPv6 or by a distributed mechanism (NBDMM). The selection of the preferred mobility anchoring solution is made by proposed decision criteria which are dynamically based on network conditions. Two decision criteria algorithms are proposed depending on the level of available information. The first one is denoted as node-assignment algorithm and it uses the network topology information to make decisions about the protocol that an AR should use. The second algorithm is called link-assignment and in addition to information about network topology, it uses BS location information in order to decide the protocol to use according to the path the MN is moving into. Additionally, we have conducted an extensive set of analytical and numerical investigations of CMM, DMM and hybrid solutions. After defining the analytical models, the expressions of mobility costs have been derived.

The results obtained show that the use of the proposed hybrid solutions outperform significantly previous mobility management schemes in terms of network resource consumption (control plane) as well as in mobility management performance (data plane). Based on the evaluations made in this work, and taking into account the flexibility of future programmable networks we conclude that significant benefits emerge from the utilization of the proposed hybrid CMM-DMM solutions in which operators would be able to handle the traffic in based on different anchoring solution depending on user traffic and topological network characteristics.

Acknowledgment

The work has been partially support by the FP7 CROSSFIRE (unCooRdinated netwOrk StrategieS for enhanced interFerence, mobIlity, radio Resource, and Energy saving management in LTE-Advanced networks) project, FP7-PEOPLE-317126.

References

[1] C. Perkins, D. Johnson, and J. Arkko, "Mobility Support in IPv6," RFC 6275, July 2011.

[2] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, "Proxy Mobile IPv6," RFC 5213, Aug. 2008.

[3] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014-2019," White Paper, Feb. 2015.

[4] D. Liu, J. C. Zuiga, P. Seite, H. Chan, and C. J. Bernardos, "Distributed Mobility Management: Current practices and gap analysis," RFC 7429, Jan. 2015.

[5] F. Giust, C. J. Bernardos, and A. De La Oliva, "Analytic Evaluation and Experimental Validation of a Network-based IPv6 Distributed Mobility Management Solution," *Mobile Computing, IEEE Transactions on*, vol. 13, no. 11, pp. 2484-2497, Nov. 2014.

[6] J. C. Zuniga, C. J. Bernardos, A. De La Oliva, T. Melia, R. Costa, and A. Reznik, "Distributed mobility management: A standards landscape," *Communications Magazine, IEEE*, vol. 51, no.3, pp. 80-87, March 2013.

[7] Dong-Hoon Shin, D. Moses, M. Venkatachalam, and S. Bagchi, "Distributed mobility management for efficient video delivery over all-IP mobile networks: Competing approaches," *Network, IEEE*, vol. 27, no. 2, pp. 28-33, April 2013.

[8] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile Data Offloading: How Much Can WiFi Deliver?," *Networking, IEEE/ACM Transactions on*, vol. 21, no. 2, pp. 536-550, April 2013.

[9] K. Ki-Sik, L. Wonjun, H. Youn-Hee, S. Myung-Ki, and Y. HeungRyeol, "Mobility management for all-IP mobile networks: Mobile IPv6 vs. Proxy Mobile IPv6," *Wireless Communications, IEEE*, vol. 15, no. 2, pp. 36-45, April 2008.

[10] Cisco 8500 Series Wireless Controller Data Sheet, Document ID 113695, February 2017. Downloadable at: `http://www.cisco.com/c/en/us/products/collateral/wireless/8500-series-wireless-controllers/data_sheet_c78-713207.html`, accessed in March 2017.

[11] H. Chan, D. Liu, P. Seite, H. Yokota, and J. Korhonen, "Requirements for Distributed Mobility Management," RFC 7333, Aug. 2014.

[12] J. H. Lee, J-M Bonnin, and X. Lagrange, "Host-based distributed mobility management support protocol for IPv6 mobile networks," in *Wireless and Mobile Computing, Networking and Communications (WiMob), 2012 IEEE 8th International Conference on*, pp. 61-68, Oct. 2012.

[13] J. H. Lee, J. M. Bonnin, P. Seite, and H. Chan, "Distributed IP mobility management from the perspective of the IETF: Motivations, Requirements, approaches, comparison, and challenges," *Wireless Communications, IEEE*, vol. 20, no. 5, pp. 159-168, Oct. 2013.

[14] H. Ali-Ahmad, M. Ouzzif, P. Bertin, and X. Lagrange, "Performance Analysis on Network-Based Distributed Mobility Management," *Wireless Personal Communications*, vol. 74, no. 4, pp. 1245-1263, Jan. 2014.

[15] T. Condeixa, and S. Sargento, "Studying the integration of distributed and dynamic schemes in the mobility management," *Computer Networks*, vol. 60, no. 26, pp. 46-59, Feb. 2014.

[16] T. Condeixa, S. Sargento, "Context-aware adaptive IP mobility anchoring," *Computer Networks*, vol. 71, no. 4, pp. 84-99, Oct. 2014.

[17] ETSI GS NFV 002, Network Functions Virtualisation (NFV); Architectural Framework, December 2014.

[18] E. Hyytia, P. Lassila, and J. Virtamo, "Spatial Node Distribution of the Random Waypoint Mobility Model with Applications," *Mobile Computing, IEEE Transactions on*, vol. 5, no. 6, pp. 680-694, June 2006.

[19] C. Bettstetter, H. Hartenstein, and X. Perez-Costa, "Stochastic Properties of the Random Waypoint Mobility Model," *Wireless Networks*, vol. 10, no. 5, pp. 555-567, Sept. 2004.

[20] M. K. Murtadha, N. K. Noordin, B. M. Ali, and F. Hashim, "Design and evaluation of distributed and dynamic mobility management approach based on PMIPv6 and MIH protocols," *Wireless Networks*, pp. 1-17, April 2015.

[21] I. Rhee, M. Shin, S. Hong, K. Lee, and Song Chong, "On the Levy-Walk Nature of Human Mobility," In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pp.13-18, April 2008.

[22] P. K. Agyapong, M. Iwamura, D. Staehle, W. Kiess and A. Benjebbour, "Design considerations for a 5G network architecture," in *IEEE Communications Magazine*, vol. 52, no. 11, pp. 65-75, Nov. 2014.